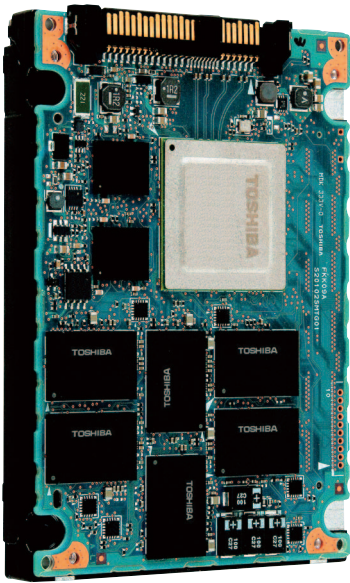
Technology of Enterprise Solid State Drive

近年、情報量とアクセス数の急速な増加に伴い、データセンタやクラウドコンピューティングで大容量かつ高性能なストレージへの需要が高まっている。ソーシャルネットワークサービス(SNS) やウェブ検索エンジンなどのインターネットサービスには多くの人がアクセスするが、それぞれが瞬時のレスポンスを期待している。また企業で運用される業務用のネットワークシステムや金融取引システムでは、マイクロ秒単位のレスポンスの違いがビジネスの成果に大きな影響を及ぼす。多くのSNSユーザのアクセスやスピードが重視されるビジネス取引を支えるために、エンタープライズサーバおよびストレージ装置のアクセス速度向上への要求がますます高まっている。この様な状況の中で、アクセス速度の速いソリッドステートドライブ(SSD) がサーバシステムのアクセス速度向上に重要な役割を果たしている。
エンタープライズ向けSSDはNAND型フラッシュメモリを記憶媒体としたストレージ装置であり、高負荷の動作条件下での、データのリード/ライト性能の超高速化と高い信頼性、高い耐久性を提供することが求められている。
システム管理者にとってSSDの採用は、処理速度や記録速度を大幅に向上できるため、非常に有効な手段である。例えば、ホストプロセッサとハードディスクドライブ(HDD)の間に位置する大容量のキャッシュとしてSSDを使用することにより、システム性能の向上が期待できる。
またSSDは省スペースかつ低消費電力といった特長も持っている。
東芝はフラッシュメモリを世界で初めて開発したメーカーとして、SSDの分野でも高い技術力を持っている。

インタフェース

サーバやストレージシステムなどでは、そのシステム構成や信頼性に応じてさまざまなインタフェースのストレージ装置が使用されている。エンタープライズ用途で最も一般的に使用されているインタフェースは、SAS(Serial Attached Small Computer System Interface (SCSI)である。SASインタフェースはプロトコルにSCSIを使用し、データ転送経路の２重化(デュアルポート)、Point-to-Point接続、全２重通信による高速データ転送、Multi-Link転送のサポートなどが特徴である。データ転送速度は、6Gbit/sから12Gbit/sへの移行が進みつつある。デュアルポートのサポートにより、ホストからの冗長パス構成をとることで高可用性を実現できる。
SATA インタフェース(Serial Advanced Technology Attachment (ATA))は、クライアントPCで一般的に使用されているが、比較的低コストで使いやすいことから、性能・機能面でシステム要件を満たす場合には、ローエンドからミッドレンジのサーバでも広く使われている。SATAインタフェースはPoint-to-Point接続、使用プロトコルはATA、転送速度は6Gbit/sなどの特徴がある。
PCIeインタフェース(Peripheral Component Interconnect (PCI) Express) を持つSSDは、高いパフォーマンスを要求されるアクセラレータやキャッシュなどのアプリケーションで使用されている。PCIe Gen3では 1レーンあたりのインタフェース速度は1GB/sでレーン数を増やすことで転送速度を上げることが可能である。また、SSDがPCIeに直接接続されるためアクセスの遅延時間が短い。プロトコルは現在のハイエンド製品では独自プロトコル、クライアント製品ではAHCI (Advanced Host Controller Interface) が使われているが、PCIe SSDの標準化プロトコルとしてNVMe (NVM Express)、SCSI over PCI-PCIe Queuing Interface (SOP-PQI)が準備されている。エンタープライズ市場ではNVMe、SOP-PQIの普及が進むと考えられる。NVMeはSSDへのデータアクセスに最適化されたプロトコルであり、SOP-PQIはSCSIプロトコルをPCIeバス上で使用する。

パフォーマンス

SSDのリード/ライト速度は、使用するNAND型フラッシュメモリ単体の性能、SSDの物理構成 (NAND型フラッシュメモリのチャネル数) およびファームウエアのアルゴリズムに依存する。SSDは、HDDのようにデータアクセスのための機械的な動作を必要とせず、電気的な動作だけでNAND型フラッシュメモリにデータのリード/ライトを行うため、高速なアクセス時間を実現している。SSDのデータ転送速度の向上には複数のチャネルを用いた並列処理が大きく寄与しており、HDDと比較して非常に速いデータ転送速度を実現できる。さらに追記型アルゴリズム、オーバプロビジョニングやUnmap/Trimコマンドなどの機能も、ランダムライト性能の向上に不可欠となっている。
NAND型フラッシュメモリの書き込み速度では、シングルレベルセル (SLC) が最速であり、次いでマルチレベルセル (MLC) である。また、SLCの方がMLCより書き換え寿命(書き換え可能回数 (write/erase cycle))で優れているため、これまでエンタープライズ向けの高速SSDには、主にSLCの NAND型フラッシュメモリが使用されてきた。しかし現在では、SSDコントローラでの書き込みアルゴリズムや並列度の最適化による性能改善、およびエラー訂正機能の強化などにより、エンタープライズ向けSSDもほとんどの用途でコストパフォーマンスに優れるMLCを使うのが主流になってきている。

エラー検出 / 訂正技術

SSDのコントローラは、データ信頼性を向上させるためにエラー訂正機能を持っている。NAND型フラッシュメモリの微細化が進むに伴い、さらに高度なエラー訂正技術が必要になってきている。エラー訂正に用いられるコードには、Hamming Code, Reed-Solomon Code, BCH (Bose-Chaudhuri-Hocquenghem) Code、LDPC (Low Density Parity Check)などがあり、NAND型フラッシュメモリの特性やSSD製品の用途に応じて用いられている。

Hamming Codeは初期のNAND型フラッシュメモリのエラー訂正で多く用いられた。演算が簡易であるため、ソフトウエアで実装することも可能である。

Reed-Solomon Codeは複数ビットを1つのシンボルとして扱い、シンボル単位でエラー訂正を行う。シンボル単位でエラー訂正を行うため、連続した誤り(バースト誤り)に強い。

BCH Codeは柔軟性があり、ブロック長や訂正能力を自由に設定でき消費電力も小さいため、現在もっとも広く使われている方式である。

LDPCは誤り訂正能力が非常に高いが消費電力、処理時間などの考慮が必要である。

東芝は、これらの得失を解決した独自のエラー訂正技術であるQSBC™を開発し、SSD製品に使用している。

パワーロス・プロテクション (Power Loss Protection)

サーバやストレージシステムなどのミッション・クリティカルなエンタープライズ向けシステムとしては、ゼロ・ダウンタイムが基本である。万が一突然電源が切断された場合にも、すでにNAND型フラッシュメモリに記録済みのデータを保持し続けることはもちろん、ホストから受信したばかりのまだデータバッファ内にあるユーザデータも、確実にNAND型フラッシュメモリに記録することが求められる。このため、SSD内にキャパシタを搭載し、このキャパシタに充電された電力を使い、NAND型フラッシュメモリへのデータ書き込みを完了させる機構を備えている。このデータ保護機能を、パワーロス・プロテクション(Power Loss Protection - PLP)と呼ぶ。

暗号

盗難やハッキングに対して安全な方法でデータを記録することは、エンタープライズサーバやストレージシステムで、もっとも重要な機能の１つである。今日では個人情報や機密情報などは、暗号化して記録することが求められている。暗号機能つきドライブ（Self-encrypting drive: SED）は、ドライブ内のAES暗号ハードウエアを使って、すべてのユーザデータを暗号化することができる。NAND型フラッシュメモリに記録されるすべてのユーザデータは、ランダムに生成された暗号鍵によって保護され、暗号鍵、パスワードやその他のセキュリティ上重要な情報も安全に管理されている。

SEDは、データ無効化機能も備えている。時間のかかる上書き作業や、ブロック単位での消去とは異なり、暗号機能を用いたデータ無効化では暗号鍵を書き変えるだけで、それまで記録されたユーザデータを瞬時に無効化できる。これにより、SSDを再利用、再配置、廃棄する際に、迅速かつ確実に記録されていたデータを無効化することが可能である。

東芝のエンタープライズ向けSSDは、Trusted Computing Group (TCG) エンタープライズSecurity Subsystem Class (SSC)準拠（SATAの場合はTCG OPAL）のデータ暗号機能を持つ。また、FIPS140-2の認証取得モデルも提供可能である。

参考文献

1）服部　正勝　他．HDD,ODD,およびSSDの技術動向．東芝レビュー．66, 8, 2011, p.30 – 35.

2）木内　英通．高性能と高信頼性を両立させたエンタープライズ向けSSD MK4001GRZB. 東芝レビュー. 66, 8, 2011, p.40 – 43.

3）木村　彰宏　他. MLC NANDフラッシュメモリを用いたエンタープライズ向け 1.6 TバイトSSD. 東芝レビュー. 68, 9, 2013, p.46 – 48.

ホームページ：**http://www.toshibastorage.com**